

+3 M A D  
1 CACCGGCGAA GGAGGATCGA ATTCCTGCAG CCCGCTATCT GCAGGCCGCC ACCATGGCCG  
GTGGCCGCTT CCTCCTAGCT TAAGGACGTC GGGCGATAGA CGTCCGGCGG TGGTACCGGC  
+3 D Y L I S G G T S Y V P D D G L T A Q Q L  
61 ACTACCTGAT TAGTGGGGGC ACGTCCTACG TGCCAGACGA CGGACTCACA GCACAGCAGC  
TGATGGACTA ATCACCCCG TGCAGGATGC ACGGTCTGCT GCCTGAGTGT CGTGTCTGTCG  
+3 L F N C G D G L T Y N D F L I L P G Y I D  
121 TCTTCAACTG CGGAGACGGC CTCACCTACA ATGACTTTCT CATTCTCCCT GGGTACATCG  
AGAAGTTGAC GCCTCTGCCG GAGTGGATGT TACTGAAAGA GTAAGAGGGA CCCATGTAGC  
+3 D F T A D Q V D L T S A L T K K I T L K T  
181 ACTTCACTGC AGACCAGGTG GACCTGACTT CTGCTCTGAC CAAGAAAATC ACTCTTAAGA  
TGAAGTGACG TCTGGTCCAC CTGGACTGAA GACGAGACTG GTTCTTTTATG TGAGAATTCT  
~~+3 T P L F S S P M D T V T E A G M A I A M A~~  
241 CCCCCTGCTT TTCCTCTCCC ATGGACACAG TCACAGAGGC TGGGATGGCC ATAGCAATGG  
GGGGTGACCA AAGGAGAGGG TACCTGTGTC AGTGTCTCCG ACCCTACCGG TATCGTTACC  
+3 A L T G G I G F I H H N C T P E F Q A N E  
301 CGCTTACAGG CGGTATTGGC TTCATCCACC ACAACTGTAC ACCTGAATTC CAGGCCAATG  
GCGAATGTCC GCCATAACCG AAGTAGGTGG TGTTGACATG TGGACTTAAG GTCCGGTTAC  
+3 E V R K V K K Y E Q G F I T D P V V L S P  
361 AAGTTCGGAA AGTGAAGAAA TAGAAGACAG GATTATCATC AGACCCTGTG GTCCTCAGCC  
TTCAAGCCTT TCACTTCTTT ATACTTGTCC CTAAGTAGTG TCTGGGACAC CAGGAGTCGG  
+3 P K D R V R D V F E A K A R H G F C G I P  
421 CCAAGGATCG CGTGCGGGAT GTTTTTGAGG CCAAGGCCCG GCATGGTTTC TGCGGTATCC  
GGTTCCCTAGC GCACGCCCTA CAAAACTCC GGTTCCGGGC CGTACCAAAG ACGCCATAGG  
+3 P I T D T G R M G S R L V G I I S S R D I  
481 CAATCACAGA CACAGGCCGG ATGGGGAGCC GCTTGGTGGG CATCATCTCC TCCAGGGACA  
GTTAGTGTCT GTGTCCGGCC TACCCCTCGG CGAACCACCC GTAGTAGAGG AGGTCCCTGT  
+3 I D F L K E E E H D C F L E E I M T K R E  
541 TTGATTTTCT CAAAGAGGAG GAACATGACT GTTTCTTGGA AGAGATAATG ACAAAGAGGG  
AACTAAAAGA GTTTCTCCTC CTTGTACTGA CAAAGAACCT TCTCTATTAC TGTTTCTCCC  
+3 E D L V V A P A G I T L K E A N E I L Q R  
601 AAGACTTGGT GGTAGCCCCT GCAGGCATCA CACTGAAGGA GGCAAATGAA ATTCTGCAGC  
TTCTGAACCA CCATCGGGGA CGTCCGTAGT GTGACTTCCT CCGTTTACTT TAAGACGTCG  
+3 R K K G K L P I V N E D D E L V A I I A  
661 GCAGCAAGAA GGGAAAGTTG CCCATTGTAA ATGAAGATGA TGAGCTTGTG GCCATCATTG  
CGTCGTTCTT CCCTTTCAAC GGGTAACATT TACTTCTACT ACTCGAACAC CGGTAGTAAC  
+3 A R T D L K K N R D Y P L A S K D A K K Q  
721 CCCGGACAGA CCTGAACAAC AATCCGGACT ACCCACTAGC CTCCAAAGAT GCCAAGAAAC  
GGGCCTGTCT GGAATTCTTC TTAGCCCTGA TGGGTGATCG GAGGTTTCTA CGGTTCTTTG  
+3 Q L L C G A A I G T H E D D K Y R L D L L  
781 AGCTGCTGTG TGGGGCAGCC ATTGGCACTC ATGAGGATGA CAAGTATAGG CTGGACTTGC  
TCGACGACAC ACCCCGTCGG TAACCGTGAG TACTCCTACT GTTCATATCC GACCTGAACG  
+3 L A Q A G V D V V V L D S S Q G N S I F Q  
841 TCGCCCAGGC TGGTGTGGAT GTAGTGGTTT TGGACTCTTC CCAGGGAAAT TCCATCTTCC  
AGCGGGTCCG ACCACACCTA CATCACAAA ACCTGAGAAG GGTCCTTTA AGGTAGAAGG  
+3 Q I N M I K Y I K D K Y P N L Q V I G G N  
901 AGATCAATAT GATCAAGTAC ATCAAAGACA AATACCCTAA TCTCCAAGTC ATTGGAGGCA  
TCTAGTTATA CTAGTTCATG TAGTTTCTGT TTATGGGATT AGAGGTTTCAG TAACCTCCGT

FIG. 1A

+3 N V V T A A Q A K N L I D A G V D A L R V  
961 ATGTGGTCAC TGCTGCCCAG GCCAAGAACC TCATTGATGC AGGTGTGGAT GCCCTGCGGG  
TACACCAGTG ACGACGGGTC CGGTTCTTGG AGTAACACG TCCACACCTA CGGGACGCCC  
+3 V G M G S G S I C I I Q E V L A C G R P Q  
1021 TGGGCATGGG AAGTGGCTCC ATCTGCATTA TCCAGGAAGT GCTGGCCTGT GGGCGGCCCC  
ACCCGTACCC TTCACCGAGG TAGACGTAAT AGGTCCTTCA CGACCGGACA CCCGCCGGGG  
+3 Q A T A V Y K V Y E Y A R R F G V P V I A  
1081 AAGCAACAGC AGTGTACAAG GTGTATGAGT ATGCACGGCG CTTTGGTGTG CCGGTCATTG  
TTCGTTGTG TCACATGTT CACATACTCA TACGTGCCGC GAAACCACAA GGCCAGTAAC  
+3 A D G G I Q N V G H I A K A L A L G A S T  
1141 CTGATGGAGG AATCCAAAAT GTGGGTCATA TTGCGAAAGC CTTGGCCCTT GGGGCCTCCA  
GACTACCTCC TTAGGTTTTA CACCAGTAT AACGCTTTCG GAACCGGAA CCCCGGAGGT  
+3 T V M M G S L L A A T T E A P G E Y F F S  
1201 CAGTCATGAT GGGCTCTCTC CTGGCTGCCA CCACTGAGGC CCCTGGTGAA TACTTCTTTT  
GTCAGTACTA CCCGAGAGAG GACCGACGGT GGTGACTCCG GGGACCACTT ATGAAGAAAA  
+3 S D G I R L K K Y R G M G S L D A M D K H  
1261 CCGATGGGAT CCGGCTAAAG AAATATCGCG GTATGGGTTC TCTCGATGCC ATGGACAAGC  
GGCTACCCTA GGCCGATTTT TTTATAGCGC CATACCCAAG AGAGCTACGG TACCTGTTCTG  
+3 H L S S Q N R Y F S E A D K I K V A Q G V  
1321 ACCTCAGCAG CCAGAACAGA TATTTCAAGT AAGCTGACAA AATCAAAGTG GCCCAGGGAG  
TGGAGTCGTC GGTCTTGTCT ATAAAGTCAC TTCGACTGTT TTAGTTTCAC CGGGTCCCTC  
+3 V S G A V Q D K G S I H K F V P Y L I A G  
1381 TGTCTGGTGC TGTGCAGGAC AAAGGGTCAA TCCACAAATT TGTCCCTTAC CTGATTGCTG  
ACAGACCAG ACACGTCCTG TTTCCCAGTT AGGTGTTTAA ACAGGGAATG GACTAACGAC  
+3 G I Q H S C Q D I G A K S L T Q V R A M M  
1441 GCATCCAAG CTCATGCCAG GACATTGGTG CCAAGAGCTT GACCCAAGTC CGAGCCATGA  
CGTAGGTTGT GAGTACGGTC CTGTAACCAC GGTTCTCGAA CTGGGTTTCAG GCTCGGTACT  
+3 M Y S G E L K F E K R T S S A Q V E G G V  
1501 TGTACTCTGG GGAGCTTAAG TTTGAGAAGA GAACGTCCTC AGCCCAGGTG GAAGGTGGCG  
ACATGAGACC CCTCGAATTC AAACCTCTCT CTTGCAGGAG TCGGGTCCAC CTTCCACCGC  
+3 V H S L H S Y E K R L F  
1561 TCCATAGCCT CCATTCGTAT GAGAAGCGGC TTTTCTGATC TAGCTCGACA TGATAAGATA  
AGGTATCGGA GGTAAGCATA CTCTTCGCCG AAAAGACTAG ATCGAGCTGT ACTATTCTAT  
1621 CATTGATGAG TTTGGACAAA CCACAAC TAGTGAATGCT TTTTGTGTA  
GTAACACTC AAACCTGTTT GGTGTTGATC TTACGTCACT TTTTTCACGA AATAAACACT  
1681 AATTTGTGAT GCTATTGCTT TATTTGTGAA ATTTGTGATG CTATTGCTTT ATTTGTAAAC  
TTAAACACTA CGATAACGAA ATAAACACTT TAAACACTAC GATAACGAAA TAAACATTGG  
1741 ATTATAAGCT GCAATAAACA AGTTAACAAC AACAATTGCA TTCATTTTAT GTTTCAGGTT  
TAATATTCTA CGTTATTGTT TCAATTGTTG TTGTTAACGT AAGTAAAATA CAAAGTCCAA  
1801 CAGGGGGAGG TGTGGGAGG TTTTAAAGC AAGTAAAACC TCTACAAATG TGGTAGATCA  
GTCCCCCTCC ACACCTCCA AAAAATTTCG TTCATTTTGG AGATGTTTAC ACCATCTAGT  
1861 TTTAAATGTT AGCGAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA  
AAATTTACAA TCGCTTCTTG TACACTCGTT TTCCGGTTCG TTTCCGGTCC TTGGCATTTT  
1921 AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCC TGACGAGCAT CACAAAAATC  
TCCGGCGCAA CGACCGCAAA AAGGTATCCG AGGCGGGGGG ACTGCTCGTA GTGTTTTTAG  
1981 GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC  
CTGCGAGTTC AGTCTCCACC GCTTTGGGCT GTCCTGATAT TTCTATGGTC CGCAAAGGGG

FIG. 1B

2041 CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG  
GACCTTCGAG GGAGCACGCG AGAGGACAAG GCTGGGACGG CGAATGGCCT ATGGACAGGC  
2101 CCTTTCTCCC TTCGGAAGC GTGGCGCTTT CTCAATGCTC ACGCTGTAGG TATCTCAGTT  
GGAAAGAGGG AAGCCCTTCG CACCGCGAAA GAGTTACGAG TCGCACATCC ATAGAGTCAA  
2161 CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCCGTT GACCCCGACC  
GCCACATCCA GCAAGCGAGG TTCGACCCGA CACACGTGCT TGGGGGGCAA GTCGGGCTGG  
2221 GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC  
CGACGCGGAA TAGGCCATTG ATAGCAGAAC TCAGGTTGGG CCATTCTGTG CTGAATAGCG  
2281 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG  
GTGACCGTCG TCGGTGACCA TTGTCTAAT CGTCTCGCTC CATACTCCG CCACGATGTC  
2341 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG AACAGTATTT GGTATCTGCG  
TCAAGAATT CACCACCGGA TTGATGCGCA TCTCATCTTC TCTCATATAA GGTATCTGCG

---

2401 CTCTGCTGAA GCCAGTTACC TTCGGAAGAAA GAGTTGGTAG CTCTTGATCC GGCAAAACAAA  
GAGACGACTT CGGTCAATGG AAGCCTTTTT CTCAACCATC GAGAACTAGG CCGTTTGTTT  
2461 CCACCGCTGG TAGCGGTGGT TTTTTGTGTT GCAAGCAGCA GATTACGCGC AGAAAAAAG  
GGTGCGGACC ATCGCCACCA AAAAAACAAA CGTTCGTCTG CTAATGCGCG TCTTTTTTTC  
2521 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT  
CTAGAGTTCT TCTAGGAAAC TAGAAAAGAT CCCCCAGACT GCGAGTCACC TTGCTTTTGA  
2581 CACGTTAAGG GATTTTGCTC ATGGCTAGTT AATTAAAGCT CAATAAACAA TCATTATTTT  
GTGCAATTCC CTAACCAG TACCGATCAA TTAATTGAC GTTATTTGTT AGTAATAAAA  
2641 CATTGGATCT GTGTGTTGGT TTTTGTGTG GGCTTGGGG AGGGGGAGGC CAGAAAGACT  
GTAACCTAGA CACACAACCA AAAACACAC CCGAACCCCT TCCCCCTCCG GTCTTACTGA  
2701 CCAAGAGCTA CAGGAAGGCA GGTCAGAGAC CCCACTGGAC AAACAGTGGC TGGACTCTGC  
GGTTCTCGAT GTCCTTCCGT CCAGTCTCTG GGGTGACCTG TTTGTCACCG ACCTGAGACG  
2761 ACCATAACAC ACAATCAACA GGGGAGTGAG CTGGATCGAG CTAGAGTCCG TTACATAACT  
TGGTATTGTG TGTTAGTTGT CCCCTCACTC GACCTAGCTC GATCTCAGGC AATGTATTGA  
2821 TACGGTAAAT GGCCCGCCTG GCTGACCGCC CAACGACCCC CGCCATTGA CGTCAATAAT  
ATGCCATTTA CCGGGCGGAC CGACTGGCGG GTTGCTGGGG GCGGGTAACT GCAGTTATTA  
2881 GACGTATGTT CCCATAGTAA CGCCAATAGG GACTTTCCAT TGACGTCAAT GGGTGGAGTA  
CTGCATACAA GGGTATCATT GCGGTTATCC CTGAAAGGTA ACTGCAGTTA CCCACCTCAT  
2941 TTTACGGTAA ACTGCCCC TGGCAGTACA TCAAGTGTAT CATATGCCAA GTACGCCCCC  
AAATGCCATT TGACGGGTGA ACCGTCACTG AGTTACATA GTATACGGT CATGCGGGGG  
3001 TATTGACGTC AATGACGGTA AATGGCCCCG CTGGCATTAT GCCCAGTACA TGACCTTATG  
ATAACTGCAG TTAATGCCAT TTACCGGGCG GACCGTAATA CGGGTCATGT ACTGGAATAC  
3061 GGACTTTCCCT ACTTGCGAGT ACATCTACGT ATTAGTCATC GCTATTACCA TGGTGATGCG  
CCTGAAAGGA TGAACCGTCA TGTAATGACA TAATCAGTAG CGATAATGGT ACCACTACGC  
3121 GTTTTGGCAG TACATCAATG GGCGTGGATA GCGGTTTGAC TCACGGGGAT TTCCAAGTCT  
CAAAACCGTC ATGTAGTTAC CCGCACCTAT CGCCAACTG AGTGCCCTA AAGGTTTCAA  
3181 CCACCCCAT TACGTCAATG GGAGTTTGT TTGGCACCAA AATCAACGGG ACTTTCCAAA  
GGTGGGGTAA CTGCAGTTAC CCTCAAACAA AACCCTGGTT TTAGTTGCCC TGAAAGGTTT  
3241 ATGTCGTAAC AACTCCGCCC CATTGACGCA AATGGGCGGT AGGCGTGTAC GGTGGGAGGT  
TACAGCATTG TTGAGGCGGG GTAACGCGT TTACCCGCCA TCCGCACATG CCACCCTCCA  
3301 CTATATAAGC AGAGCTCGTT TAGTGAACCG TCAGATCGCC TGGAGACGCC ATCCACGCTG  
GATATATTCT TCTCGAGCAA ATCACTTGGC AGTCTAGCGG ACCTCTGCGG TAGGTGCGAC  
3361 TTTTGACCTC CATAGAAGAC ACCGGGACCG ATCCAGCCTC CGCGGCCGGG AACGGTGCAT  
AAAACCTGGAG GTATCTTCTG TGGCCTGGC TAGGTCGGAG GCGCCGGCCC TTGCCACGTA

FIG. 1C

3421 TGGAACGCGG ATTCCCCGTG CCAAGAGTGA CGTAAGTACC GCCTATAGAG TCTATAGGCC  
ACCTTGCGCC TAAGGGGCAC GGTCTCACT GCATTCATGG CGGATATCTC AGATATCCGG  
3481 CACCCCCTTG GCTTCTTATG CATGCTATAC TGTTTTTGGC TTGGGGTCTA TACACCCCGG  
GTGGGGGAAC CGAAGAATAC GTACGATATG AAAAAACCG AACCCAGAT ATGTGGGGGC  
3541 CTTCCCTCATG TTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA  
GAAGGAGTAC AATATCCACT ACCATATCGA ATCGGATATC CACACCAAT AACTGGTAAT  
3601 TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG  
AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC  
3661 CCACAACCTC CTTTATTGGC TATATGCCAA TACACTGTCC TTCAGAGACT GACACGGACT  
GGTGTGAGA GAAATAACCG ATATACGGTT ATGTGACAGG AAGTCTCTGA CTGTGCCTGA  
3721 CTGTATTTTT ACAGGATGGG GTCTCATTTA TTATTTACAA ATTCACATAT ACAACACCAC  
GACATAAAAA TGTCCTACCC CAGAGTAAAT AATAAATGTT TAAGTGTATA TGTTGTGGTG  
3781 CGTCCCCAGT GCCCGCAGTT TTTATTAAAC ATAAGCTCCG ATCTCCACCC CAATCTCCCG  
GCAGGGGTCA CGGGCGTCAA AAATAATTTG TATTGCACCC TAGAGGTGCG CTTAGAGCCC  
3841 TACGTGTTCC GGACATGGGC TCTTCTCCGG TAGCGGCGGA GCTTCTACAT CCGAGCCCTG  
ATGCACAAGG CCTGTACCCG AGAAGAGGCC ATCGCCGCC TCGAAGATGTA GGCTCGGGAC  
3901 CTCCCATGCC TCCAGCGACT CATGGTCGCT CGGCAGCTCC TTGCTCCTAA CAGTGGAGGC  
GAGGGTACGG AGGTCGCTGA GTACCAGCA GCGTCGAGG AACGAGGATT GTCACCTCCG  
3961 CAGACTTAGG CACAGCAGA TGCCACCCAC CACCACTGTG CCGACAAGG ACGTGGCGGT  
GTCTGAATCC GTGTCTGTCT ACGGGTGGTG GTGGTCACAC GGCGTGTTC GGCACGCCCA  
4021 AGGGTATGTG TCTGAAAATG AGCTCGGGGA GCGGGCTTGC ACCGCTGACG CATTTGGAAG  
TCCCATACAC AGACTTTTAC TCGAGCCCCT CGCCCGAACG TGGCGACTGC GTAAACCTTC  
4081 ACTTAAGGCA GCGGCAGAAG AAGATGCAGG CAGCTGAGTT GTTGTGTTCT GATAAGAGTC  
TGAATTCCGT CGCCGTCTTC TTCTACGTCC GTCGACTCAA CAACACAAGA CTATTCTCAG  
4141 AGAGGTAACCT CCCGTTGCGG TGCTGTTAAC GGTGGAGGGC AGTGTAGTCT GAGCAGTACT  
TCTCCATTGA GGGCAACGCC ACGACAATTG CCACCTCCCG TCACATCAGA CTCGTCATGA  
4201 CGTTGCTGCC GCGCGCGCCA CCAGACATAA TAGCTGACAG ACTAACAGAC TGTTCTTTTC  
GCAACGACGG CCGCGCGCGT GGTCTGTATT ATCGACTGTC TGATTGTCTG ACAAGGAAAG  
MCS  
\*\*\*\*\*  
4261 CATGGGTCTT TTCTGCAGTC ACCCGGGGGA TCCTTCGAAC GTAGCTCTAG ATTGAGTCGA  
GTACCCAGAA AAGACGTCAG TGGGCCCCCT AGGAAGCTTG CATCGAGATC TAACTCAGCT  
4321 CGTTACTGGC CGAAGCCGCT TGGAATAAG CCGGTGTGCG TTTGTCTATA TGTTATTTTC  
GCAATGACCG GCTTCGGCGA ACCTTATTCC GGCCACACGC AAACAGATAT ACAATAAAAG  
4381 CACCATATTG CCGTCTTTTG GCAATGTGAG GGCCCGGAAA CCTGGCCCTG TCTTCTTGAC  
GTGGTATAAC GGCAGAAAAC CGTTACACTC CCGGGCCTTT GGACCGGGAC AGAAGAAGTG  
4441 GAGCATTCCT AGGGGTCTTT CCCCTCTCGC CAAAGGAATG CAAGGTCTGT TGAATGTCGT  
CTCGTAAGGA TCCCAGAAA GGGGAGAGCG GTTTCCTTAC GTTCCAGACA ACTTACAGCA  
4501 GAAGGAAGCA GTTCCTCTGG AAGCTTCTTG AAGACAAACA ACGTCTGTAG CGACCCTTTG  
CTTCCTTCGT CAAGGAGACC TTCGAAGAAC TTCTGTTTGT TGCAGACATC GCTGGGAAAC  
4561 CAGGCAGCGG AACCCCCAC CTGGCGACAG GTGCCTCTGC GGCCAAAAGC CACGTGTATA  
GTCCGTCGCC TTGGGGGGTG GACCGCTGTC CACGGAGACG CCGGTTTTTC GTGCACATAT  
4621 AGATACACCT GCAAAGGCGG CACAACCCCA GTGCCACGTT GTGAGTTGGA TAGTTGTGGA  
TCTATGTGGA CGTTTCCGCC GTGTTGGGGT CACGGTGCAA CACTCAACCT ATCAACACCT  
4681 AAGAGTCAAA TGGCTCTCCT CAAGCGTATT CAACAAGGGG CTGAAGGATG CCCAGAAGGT  
TTCTCAGTTT ACCGAGAGGA GTTCGCATTA GTTGTTCCTC GACTTCCTAC GGTCTCTCCA  
4741 ACCCCATTGT ATGGGATCTG ATCTGGGGCC TCGGTGCACA TGCTTTACAT TGTTTATGTC  
TGGGGTAACA TACCCTAGAC TAGACCCCGG AGCCACGTGT ACGAAATGTA CACAAATCAG

FIG. 1D

4801 GAGGTTAAAA AAACGTCTAG GCCCCCGGAA CCACGGGGAC GTGGTTTTCC TTTGAAAAAC  
CTCCAATTTT TTTGCAGATC CGGGGGGCTT GGTGCCCCCTG CACCAAAAAGG AAACTTTTTG  
4861 ACGATAATAC CATGGGTAAC TGATATCTAC TAGTTGTGAC CGGCGCCTAG TGTGACAAT  
TGCTATTATG GTACCCATTC ATCATAGATG ATCAACACTG GCCGCGGATC ACAACTGTTA  
4921 TAATCATCGG CATAGTATAT CGGCATAGTA TAATACGACT CACTATAGGA GGGCCACCAT  
ATTAGTAGCC GTATCATATA GCCGTATCAT ATTATGCTGA GTGATATCCT CCCGGTGGTA  
4981 GTCGACTACT AACCTTCTTC TCTTTCCTAC AGCTGAGATC ACCGGTAGGA GGGCCATCAT  
CAGCTGATGA TTGGAAGAAG AGAAAGGATG TCGACTCTAG TGGCCATCCT CCCGGTAGTA  
5041 GAAAAAGCCT GAACTACCG CGACGTCTGT CGCGAAGTTT CTGATCGAAA AGTTCGACAG  
CTTTTTTCGGA CTTGAGTGGC GCTGCAGACA GCGCTTCAAA GACTAGCTTT TCAAGCTGTC  
5101 CGTCTCCGAC CTGATGCAGC TCTCGGAGGG CGAAGAATCT CGTGCTTTCA GCTTCGATGT  
~~GCAGAGGCTC CACTACCTCG AGAGGCTCCG GCTTCTTACG GCACGAAAGT CGAAGCTACA~~  
5161 AGGAGGGCGT GGATATGTCC TGCGGGTAAA TAGCTGCGCC GATGGTTTCT ACAAAGATCG  
TCCTCCCGCA CCTATACAGG ACGCCCATTT ATCGACGCGG CTACCAAAGA TGTTCCTAGC  
5221 TTATGTTTAT CGGCACTTTG CATCGGCCGC GCTCCCGATT CCGGAAAGTGC TTGACATTGG  
AATACAAATA GCCGTGAAAC GTAGCCGGCG CGAGGGCTAA GGCTTCACG AACTGTAACC  
5281 GGAATTCAGC GAGAGCTGA CCTATTGCAT CTCCCGCCGT GCACAGGGTG TCACGTTGCA  
CCTTAAGTCG CTCTCGGACT GGATAACGTA GAGGGCGGCA CGTGTCCAC AGTGCAACGT  
5341 AGACCTGCCT GAAACCGAAC TGCCCGCTGT TCTGCAACCC GTCGCGGAG TCATGGATGC  
TCTGGACGGA CTTTGGCTTG ACGGGCGACA AGACGTTGGG CAGCGCCTCG AGTACCTACG  
5401 GATCGCTGCG GCCGATCTTA GCCAGACGAG CGGGTTCGGC CCATTGCGAC CGCAAGGAAT  
CTAGCGACGC CGGCTAGAAT CGGTCTGCTC GCCCAAGCCG GGTAAGCCTG GCGTTCCTTA  
5461 CGGTCAATAC ACTACATGGC GTGATTTTCAT ATGCGCGATT GCTGATCCCC ATGTGTATCA  
GCCAGTTATG TGATGTACCG CACTAAAGTA TACGCGCTAA CGACTAGGGG TACACATAGT  
5521 CTGGCAAACGT GTGATGGACG ACACCGTCAG TGCGTCCGTC GCGCAGGCTC TCGATGAGCT  
GACCGTTTGA CACTACCTGC TGTGGCAGTC ACGCAGGCAG CGCGTCCGAG AGCTACTCGA  
5581 GATGCTTTGG GCCGAGGACT GCCCCGAAGT CCGGCACCTC GTGCACGCGG ATTTCCGGCTC  
CTACGAAACC CGGCTCCTGA CGGGGCTTCA GGCCGTGGAG CACGTGCGCC TAAAGCCGAG  
5641 CAACAATGTC CTGACGGACA ATGGCCGCAT AACAGCGGTC ATTGACTGGA GCGAGGCGAT  
GTTGTTACAG GACTGCCTGT TACCGGCGTA TTGTCGCCAG TAACTGACCT CGTCCGCTA  
5701 GTTCGGGGAT TCCCAATACG AGGTCGCCAA CATCTTCTTC TGGAGGCCGT GGTGCGCTTG  
CAAGCCCCTA AGGGTTATGC TCCAGCGGTT GTAGAAGAAC ACCTCCGGCA CCAACCGAAC  
5761 TATGGAGCAG CAGACGCGCT ACTTCGAGCG GAGGCATCCG GAGCTTGCAG GATCGCCGCG  
ATACCTCGTC GTCTGCGCGA TGAAGCTCGC CTCCGTAGGC CTCGAACGTC CTAGCGGCGC  
5821 GCTCCGGGCG TATATGCTCC GCATTGGTCT TGACCAACTC TATCAGAGCT TGGTTGACGG  
CGAGGCCCGC ATATACGAGG CGTAACCAGA ACTGGTTGAG ATAGTCTCGA ACCAACTGCC  
5881 CAATTTCGAT GATGCAGCTT GGGCGCAGGG TCGATGCGAC GCAATCGTCC GATCCGGAGC  
GTTAAAGCTA CTACGTCGAA CCCGCGTCCC AGCTACGCTG CGTTAGCAGG CTAGGCCTCG  
5941 CGGGACTGTC GGGCGTACAC AAATCGCCCC CAGAAGCGCG GCCGTCTGGA CCGATGGCTG  
GCCCTGACAG CCCGCATGTG TTTAGCGGGC GTCTTCGCGC CGGCAGACCT GGCTACCGAC  
6001 TGTAGAAGTC GCGTCTGCGT TCGACCAGGC TGC GCGTTCCT CGCGGCCATA GCAACCGACG  
ACATCTTCAG CGCAGACGCA AGCTGGTCCG ACGCGCAAGA GCGCCGGTAT CGTTGGCTGC  
6061 TACGGCGTTG CGCCCTCGCC GGCAGCAAGA AGCCACGGAA GTCCGCCCGG AGCAGAAAAT  
ATGCCGCAAC GCGGGAGCGG CCGTCGTTCT TCGGTGCCTT CAGGCGGGCC TCGTCTTTTA  
6121 GCCCACGCTA CTGCGGGTTT ATATAGACGG TCCCCACGGG ATGGGGAAAA CCACCACCAC  
CGGGTGCGAT GACGCCCAAA TATATCTGCC AGGGGTGCC TACCCCTTTT GGTGGTGGTG

FIG. 1E

6181 GCAACTGCTG GTGGCCCTGG GTTCGCGCGA CGATATCGTC TACGTACCCG AGCCGATGAC  
CGTTGACGAC CACCGGGACC CAAGCGCGCT GCTATAGCAG ATGCATGGGC TCGGCTACTG  
6241 TTACTGGCGG GTGCTGGGGG CTTCCGAGAC AATCGCGAAC ATCTACACCA CACAACACCG  
AATGACCGCC CACGACCCCC GAAGGCTCTG TTAGCGCTTG TAGATGTGGT GTGTTGTGGC  
6301 CCTCGACCAG GGTGAGATAT CGGCCGGGGA CGCGGCGGTG GTAATGACAA GCGCCCAGAT  
GGAGCTGGTC CCACTCTATA GCCCGCCCCT GCGCCGCCAC CATTACTGTT CGCGGGTCTA  
6361 AACAAATGGC ATGCCTTATG CCGTGACCGA CGCCGTTCTG GCTCCTCATA TCGGGGGGGA  
TTGTTACCCG TACGGAATAC GGCACGTGGT GCGGCAAGAC CGAGGAGTAT AGCCCCCCT  
6421 GGCTGGGAGC TCACATGCCC CGCCCCCGGC CCTCACCCCTC ATCTTCGACC GCCATCCCAT  
CCGACCTCG AGTGTACGGG GCGGGGGCCG GGAGTGGGAG TAGAAGCTGG CGGTAGGGTA  
6481 CGCCGCCCTC CTGTGCTACC CGGCCGCGCG GTACCTTATG GGCAGCATGA CCCCCAGGC  
~~CGCCGCCCTC CTGTGCTACC CGGCCGCGCG GTACCTTATG GGCAGCATGA CCCCCAGGC~~  
6541 CGTGCTGGCG TTCGTGGCCC TCATCCCGCC GACCTTGCCC GGCACCAACA TCGTGCTTGG  
GCACGACCGC AAGCACCGGG AGTAGGGCGG CTGGAACGGG CCGTGTTGT AGCACGAACC  
6601 GGCCCTTCCG GAGGACAGAC ACATCGACCG CCTGGCCAAA CGCCAGCGCC CCGCGAGCG  
CCGGGAAGGC CTCCTGTCTG TGTAGCTGGC GGACCGGTTT GCGGTGCGCG GGCCGCTCGC  
6661 GCTGGACCTG GCTATGCTGG CTGCGATTG CCGGTTTAC GGGCTACTTG CCAATACGGT  
CGACCTGGAC CGATACGACC GACGCTAAGC GCGCAAATG CCCGATGAAC GGTATGCCA  
6721 GCGGTATCTG CAGTGCGGCG GGTGCTGGCG GGAGGACTGG GGACAGCTTT CGGGGACGGC  
CGCCATAGAC GTCACGCCGC CCAGACCGC CCTCCTGACC CTGTGCGAAA GCCCTGCCG  
6781 CGTGCCGCCC CAGGGTGCCG AGCCCCAGAG CAACGCGGGC CCACGACCCC ATACGGGGGA  
GCACGGCGGG GTCCCACGGC TCGGGGCTC GTTGCGCCG GGTGCTGGG TATAGCCCT  
6841 CACGTTATTT ACCCTGTTTC GGGCCCCCGA GTTGCTGGCC CCAACGGCG ACCTGTATAA  
GTGCAATAAA TGGGACAAAG CCCGGGGGCT CAACGACCG GGGTTGCCG TGGACATATT  
6901 CGTGTGTTGCC TGGGCCTTGG ACGTCTTGG CAAACGCCTC CGTTCCATGC ACGTCTTTAT  
GCACAAACGG ACCCGGAACC TGCAGAACCG GTTTGCGGAG GCAAGGTACG TGCAGAAATA  
6961 CCTGGATTAC GACCAATCGC CCGCCGGCTG CCGGGACGCC CTGCTGCAAC TTACCTCCGG  
GGACCTAATG CTGGTTAGCG GCGGCGCGAC GGCCCTGCGG GACGACGTTG AATGGAGGCC  
7021 GATGGTCCAG ACCCACGTCA CCACCCCGG CTCCATACCG ACGATATGCG ACCTGGCGCG  
CTACCAGGTC TGGGTGCAGT GGTGGGGGCC GAGGTATGGC TGCTATACGC TGGACCGCG  
7081 CACGTTTGCC CGGGAGATGG GGGAGGCTAA CTGAGTCGAG AATTTCGCTAG AGGGCCCTAT  
GTGCAACCG GCCCTCTACC CCTCCGATT GACTCAGCTC TTAAGCGATC TCCCGGGATA  
7141 TCTATAGTGT CACCTAAATG CTAGAGCTCG CTGATCAGCC TCGACTGTGC CTCTAGTTG  
AGATATCACA GTGGATTTAC GATCTCGAGC GACTAGTCGG AGCTGACACG GAAGATCAAC  
7201 CCAGCCATCT GTTGTGTTGCC CCTCCCCGT GCCTTCCTTG ACCCTGGAAG GTGCCACTCC  
GGTCGGTAGA CAACAAACGG GGAGGGGGCA CGGAAGGAAC TGGGACCTTC CACGGTGAGG  
7261 CACTGTCCTT TCCTAATAAA ATGAGGAAAT TGCATCGCAT TGTCTGAGTA GGTGTCATT  
GTGACAGGAA AGGATTATTT TACTCCTTTA ACGTAGCGTA ACAGACTCAT CCACAGTAAG  
7321 TATTCTGGGG GGTGGGGTGG GGCAGGACAG CAAGGGGGAG GATTGGGAAG ACAATAGCAG  
ATAAGACCCC CCACCCCACC CCGTCTGTC GTTCCCCCTC CTAACCTTC TGTATCGTC  
7381 GCATGCGCAG GGCCCAATTG CTCGAGCGGC CGCAATAAAA TATCTTTATT TTCATTACAT  
CGTACGCGTC CCGGGTTAAC GAGCTCGCCG GCGTTATTTT ATAGAAATAA AAGTAATGTA  
7441 CTGTGTGTTG GTTTTTTGTG TGAATCGTAA CTAACATACG CTCTCCATCA AAACAAAACG  
GACACACAAC CAAAAACAC ACTTAGCATT GATTGTATGC GAGAGGTAGT TTTGTTTTGC  
7501 AAACAAAACA AACTAGCAAA ATAGGCTGTC CCCAGTGCAA GTGCAGGTGC CAGAACATTT  
TTTGTGTTGT TTGATCGTTT TATCCGACAG GGGTCACGTT CACGTCCACG GTCTTGTA

FIG. 1F

7561 CTCTATCGAA GGATCTGCCA TCGCTCCGGT GCCCGTCAGT GGGCAGAGCG CACATCGCCC  
 GAGATAGCTT CCTAGACGCT AGCGAGGCCA CGGGCAGTCA CCCGTCTCGC GTGTAGCGGG  
 7621 ACAGTCCCCG AGAAGTTGGG GGGAGGGGTC GGCAATTGAA CCGGTGCCTA GAGAAGGTGG  
 TGTCAGGGGC TCTTCAACCC CCCTCCCCAG CCGTTAACTT GGCCACGGAT CTCTTCCACC  
 7681 CGCGGGGTAA ACTGGGAAAG TGATGTCGTG TACTGGCTCC GCCTTTTTCC CGAGGGTGGG  
 GCGCCCCATT TGACCCTTTC ACTACAGCAC ATGACCGAGG CGGAAAAGG GCTCCCACCC  
 7741 GGAGAACCGT ATATAAGTGC AGTAGTCGCC GTGAACGTTT TTTTTCGCAA CGGGTTTGCC  
 CCTCTTGGCA TATATTCACG TCATCAGCGG CACTTGCAAG AAAAAGCGTT GCCCAAACGG  
 7801 GCCAGAACAC AGCTGAAGCT TCGAGGGGCT CGCATCTCTC CTTACGCGC CCGCCGCCCT  
 CGGTCTTGTG TCGACTTCGA AGCTCCCCGA GCGTAGAGAG GAAGTGCGCG GCGGCGGGA  
 7861 ACCTGAGGCC GCCATCCACG CCGGTTGAGT CCGGTTCTGC CGCCTCCCGC CTGTGGTGCC  
 TGGACTCCGG CGGTAGGTGC GGCCAACTCA GCGCAAGACG GCGGAGGGCG GACACCACGG  
 7921 TCCTGAACTG CGTCCGCCGT CTAGGTAAGT TTAAAGCTCA GGTCGAGACC GGGCCTTTGT  
 AGGACTTGAC GCAGGCGGCA GATCCATTCA AATTTTCGAGT CCAGCTCTGG CCCGGAAAACA  
 7981 CCGGCGCTCC CTTGGAGCCT ACCTAGACTC AGCCGGCTCT CCACGCTTTG CCTGACCCTG  
 GGCCGCGAGG GAACCTCGGA TGGATCTGAG TCGGCCGAGA GGTGCGAAAC GGAAGTGGGAC  
 8041 CTTGCTCAAC TCTACGTCTT TGTTTCGTTT TCTGTTCTGC GCCGTTACAG ATCCAAGCTG  
 GAACGAGTTG AGATGCAGAA ACAAAGCAAA AGACAAGACG CGGCAATGTC TAGGTTTCGAC  
 8101 TGACCGGCGC CTACGTAAGT GATATCTACT AGATTATCA AAAAGAGTGT TGACTTCTGA  
 ACTGGCCGCG GATGCATTCA CTATAGATGA TCTAAATAGT TTTTCTCACA ACTGAACACT  
 8161 GCGCTCACAA TTGATACTTA GATTCATCGA GAGGGACACG TCGACTACTA ACCTTCTTCT  
 CGCGAGTGTT AACTATGAAT CTAAGTAGCT CTCCCTGTGC AGCTGATGAT TGGAAGAAGA  
 8221 CTTTCCTACA GCTGAGAT  
 GAAAGGATGT CGACTCTA

FIG. 1G